

Application no/date: 1983-130677 [1983/ 7/ 1]
 Date of request for examination: [1987/ 5/ 1]
 Accelerated examination ()
 Public disclosure no/date: 1985- 22193 [Translates] [1985/ 2/ 4]
 Examined publication no/date (old law): []
 Registration no/date: []
 Examined publication date (present law): []
 PCT application no:
 PCT publication no/date: []
 Applicant: NEC CORP
 Inventor: TSUKAMOTO NOBORU
 IPC: G10L 3/00
 FI: G10L 3/00, 571K G10L 3/02, 301 G10L 3/02, 301C
 G10L 3/02, 301E

F-Term: 5D015CC00, CC17, CC18, DD01, DD03, EE05

Expanded classification: 425, 453

Fixed keyword: RIQS

Citation: [.] (, , ,)

Title of invention: VOICE RECOGNITION EQUIPMENT

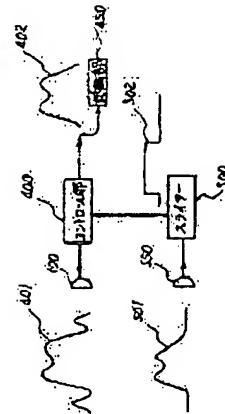
Abstract: Voice interval of the language which uttered by using microphone detecting jolt to produce in attendant tsute human organism for utterance other than microphone detecting air jolt by utterance together is arrested precisely, and high acknowledgement performance is got.

Input signal 401 of setsuhanashikata microphone 100 is input into input signal control department 400.

Pickup sign 501 which did is sliced in slicer department 500 than throat microphone 550 by suretsushiyorudoreberu decided beforehand, and it is output audio signal from human organism for sign 502.

This sign 502 is employed as the gate signal which gate control makes input signal from setsuhanashikata microphone 100.

Sign 502 is output in department 450 that input signal 401 recognizes only at the time of H level.



Best Available Copy

⑨ 日本国特許庁 (JP)
 ⑩ 公開特許公報 (A)

⑪ 特許出願公開
 昭60-22193

⑫ Int. Cl.
 G 10 L 3/00

識別記号 厅内整理番号
 7350-5D

⑬ 公開 昭和60年(1985)2月4日
 発明の数 1
 審査請求 未請求

(全 3 頁)

④ 音声認識装置

⑤ 特 願 昭58-130677
 ⑥ 出 願 昭58(1983)7月18日
 ⑦ 発明者 塚本界

東京都港区芝五丁目33番1号日本電気株式会社内
 ⑧ 出願人 日本電気株式会社
 東京都港区芝5丁目33番1号
 ⑨ 代理人 弁理士 内原晋

明細書

1. 発明の名称

音声認識装置

2. 特許請求の範囲

認識されるべき音声を入力する第一のマイクロホンと、音声を发声する发声源の振動を取らえる第二のマイクロホンと、前記第二のマイクロホンの出力に応じて前記第一のマイクロホンの出力をゲートする制御部とを具備し、前記制御部の出力を受けて音声認識を行うことを特徴とする音声認識装置。

3. 発明の詳細な説明

本発明は、発声された言葉を認識する音声認識装置に関するものである。

近年、コンピュータや各種制御装置等における入力装置として音声認識装置が本格的実用化を向かえるに至っている。人間の話す言葉をそのまま

認識できる音声認識装置は、利用のための特徴を挙げてもいらず、視線や手足が拘束されないなど多くの利点があることは既に周知の通りである。しかししながら、このより本利点の反面、装置を使用する周囲騒音により、認識性能が大きく影響を受けるという欠点があることも事実である。特にランダムに発生する衝撃音は認識装置に及ぼす性能低下の影響が大きく、問題となっている。

音声認識装置では、周囲騒音の影響を少なくする目的でノイズキャンセリング型のマイクロホンを使用しているのが一般的であり、ノイズキャンセリング型のマイクロホンを使用することにより周囲騒音のレベルは減少し、有効であるが、音声レベルの小さな領域では、なおかつ周囲騒音の影響を受けてしまう。すなわちこの音声レベルの小さな領域とは、特に発声された音声の始端、終端部分に存在し、音声区間を検出することに対し大きな障害となっていた。又、騒音に影響を受けない、スロートマイクや、イヤーマイクの使用も考えられるがこれらのマイクロホンの場合音声の高

Best Available Copy

特開昭60- 22193(2)

い周波数部分の有効が得られにくくことから認識率が、静かな所で使用した通常のマイクロホンに比較して明らかに低く、この方法も、又、問題の解決には至っていない。

周囲駆音は、定常的なものと、非定常なものとに分けることができる。定常的なものは、その駆音が 85 dB 仰位までならば、認識性能にはほとんど影響を受けない。これは、あらかじめ发声していない時に周波数音のレベルを把握できることによる。それに比べ非定常なものは、发声する駆音の時期、大きさが分らず、この種の駆音の影響を受けないで認識するということとは難しい。

本発明の目的は、发声した音楽の音声区间を外部駆音に影響されることなく、安定して正確に捉えることができ、安定した高い認識性能が得られる音声認識装置を提供することである。

本発明によれば、音声認識のために使用している駆声による空気振動をピックアップするマイクロホンの他に发声に伴って人体に生じる振動をピックアップするスロートマイクあるいはイヤーマ

に小さな駆音となっている。従って周囲駆音との関係を見ると始端、終端付近に駆音が入る事がレベル的に最も影響を受ける。又、時間方向上から見ても始端、終端が“ス”して音声区间が実質のものより長くなり、发声した音楽のパターンと異ったものとして捉えられることになる。これに比べ、始端、終端以外の部分は、駆音が混入してもその影響は比較的少ない。このことから周囲駆音に関しては、音声区间の始端、終端を正しく検出できれば認識性能の改善に大きく寄与することがわかる。

図 3 は本発明の一実施例である。一般の音声認識装置で使用されている長話型マイクロホン 100 の入力信号 401 が入力信号コントロール部 400 に入力される。この信号は図のように波形の前後にある山形の部分が周囲駆音をひろった信号である。又、人体からの音声信号をスロートマイク 550 よりピックアップしたものが信号 501 であり、周囲駆音の影響は全くないものである。信号 501 は、あらかじめ決められたスレッショル

イク等を用い認識部に入力する音声入力を限定する音声認識装置が得られる。

以下、本発明を図を用いて詳細に説明する。図 1 は、実用化されている音声認識装置の信号入力部のブロック図である。オペレータにより操作された音楽は、空気振動に変換され空気中を伝播し、マイクロホン 100 に入力され周波数分析部 200 でサンプリングされいくつかの周波数チャンネルごとにデジタル化され、音声検出部 300 で入力された信号が音声か否か分析される。

次に音声検出のメカニズム例を図 2 により説明する。一般に入力された信号は、そのエネルギーのレベル値により音声の検出を行っている。入力信号レベルが TH 1 を超え以降の時間 t の積分値が TH 2 を超えている時に音声の始端検出としている。同様に音声の終端検出は、入力信号レベルが TH 3 以下になり、なおかつ TH 3 以下になつた時点より時間 t の間連続して TH 3 以下の状態であったとき行っている。この始端、終端レベルは音声区间中の信号レベルに比較すると相当

トレベル化よりスライサー部 510 でスライスされ信号 502 として出力される。この信号 502 は、接話型マイクロホンからの入力信号をゲートコントロールするゲート信号として使用される。信号 502 が “H” レベルの時の入力信号 401 が認識部 450 に出力されることになる。402 はこの出力信号であり、入力信号 401 での周囲駆音の信号部分が取り除かれたものである。このように従来のマイクロホンからの入力信号の他に人体からの音声信号を音声区间検出用として使用することにより、周囲駆音に全く影響されない音声認識装置を実現することが可能である。

本実施例は、人体からの音声信号を音声区间検出用として特に始端、終端部の影響を除去する目的のみ使用しているが、さらに音声中の駆音の影響を除去する目的から接話型マイクロホンからの信号と人体からの音声信号のエネルギー値の相関をとるなどの手法で、さらに高認識性能の音声認識装置を実現することも可能である。

Best Available Copy

特許昭60- 22193(3)

4. 図面の簡単な説明

図1は、現状の音声認識装置の信号入力部のブロック図例である。図2は、音声検出のメカニズムの一例を表わす音声入力波形である。図3は、本発明の一実施例である。

代理人弁理士 内原 喬

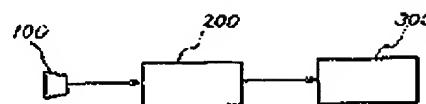


図 1

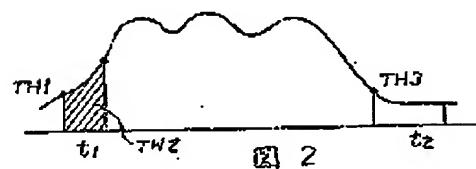


図 2

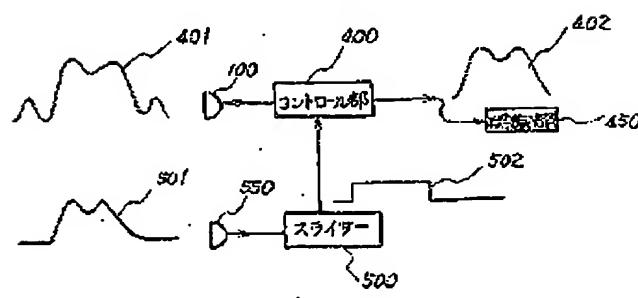


図 3

Best Available Copy